Recovering developmental bivariate trajectories in accelerated longitudinal designs with dynamic continuous time modeling

Nuria Real-Brioso[1],

Eduardo Estrada[1],

Pablo F. Cáncer[2]

[1]*Department of Social Psychology and Methodology, Universidad Autónoma de Madrid, Spain.*

[2]*UNINPSI Clinical Psychology Center, Department of Psychology, Universidad Pontificia Comillas, Madrid, Spain*

This manuscript has been peer-reviewed and accepted for publication in the Taylor & Francis journal *Structural Equation Modeling: A multidisciplinary journal.* This paper is not the copy of the record and may not exactly replicate the final authoritative version of the article. The final article will be available, upon publication, vita its DOI: 10.1080/10705511.2023.2277651

To cite this article:

Real-Brioso, N., Estrada, E., Cáncer, P.F. (2023). Recovering Developmental Bivariate Trajectories in Accelerated Longitudinal Designs with Dynamic Continuous Time Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 1-16. https://doi.org/10.1080/10705511.2023.2277651

Author Note

*Correspondence should be sent to:

Eduardo Estrada, Ph.D., eduardo.estrada.rs@gmail.com

Dpt. of Social Psychology and Methodology. Universidad Autónoma de Madrid

Ivan Pavlov 6. 28049, Madrid (Spain)

ORCID: Nuria Real-Brioso (0000-0002-3890-5062), Eduardo Estrada (0000-0003-0899-4057) and Pablo F. Cáncer (0000-0001-9279-8440).

Acknowledgements: This work was funded by the Ministry of Science and Innovation of Spain (ref. PID2019-107570GA-I00 / AEI / doi: 10.13039/501100011033), granted to EE.

Abstract

Accelerated longitudinal designs (ALDs) provide an opportunity to capture long developmental periods in a shorter time framework using a relatively small number of assessments. Prior literature has investigated whether univariate developmental processes can be characterized with data obtained from ALDs. However, many important questions in psychology and related sciences imply working with several variables that are intercorrelated as they unfold over time, such as cognitive and cortical development. Therefore, bivariate developmental models are required. This study aimed to assess the effectiveness of continuous-time bivariate Latent Change Score (CT-BLCS) models for recovering the trajectories of two interdependent developmental processes using data from diverse ALDs. Through a Monte Carlo simulation study, the efficacy of different sampling designs and sample sizes was examined. The study fills a gap in the literature by examining the performance of ALDs in bivariate systems, providing specific recommendations for future application of ALDs for studying interrelated developmental variables.

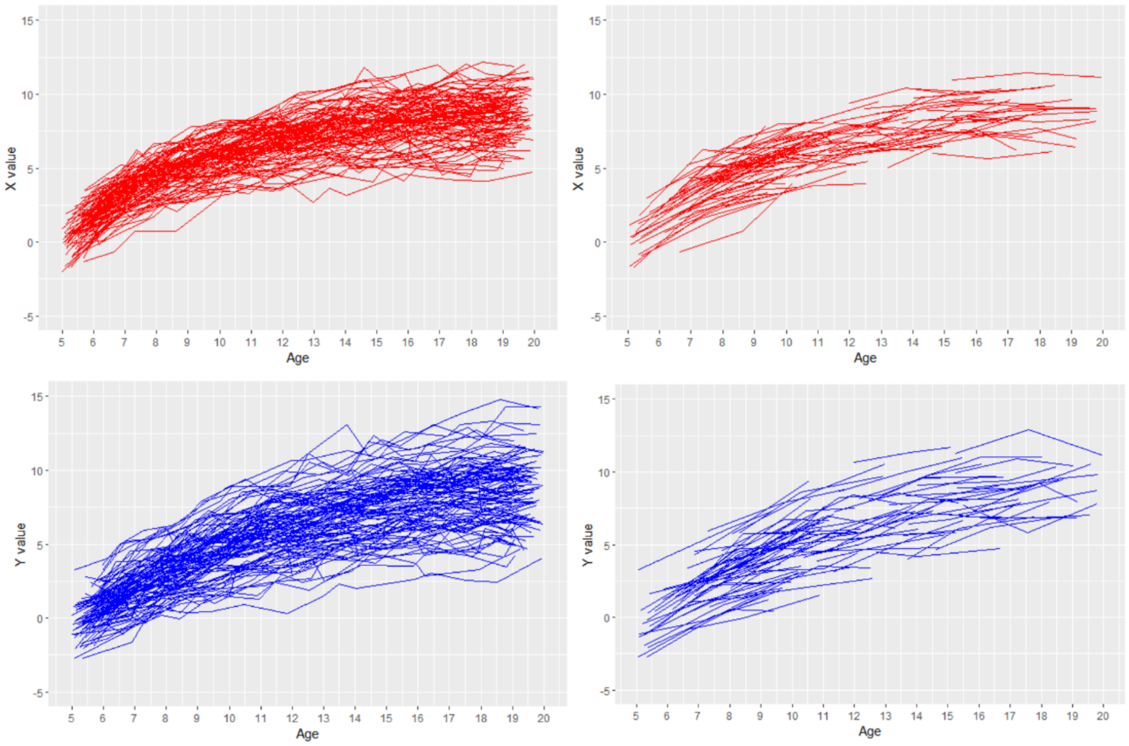
Keywords: Accelerated longitudinal design, bivariate latent change score, continuous time modeling, state-space models, bivariate developmental process

Recovering developmental bivariate trajectories in accelerated longitudinal designs with dynamic continuous time modeling

Accelerated longitudinal designs (ALDs; Bell, 1953, 1954; Duncan et al., 1996) also referred to as cohort-sequential (Nesselroade & Baltes, 1979), overlapping cohort designs (Oud, 2001) or cross-sequential designs (Schaie, 1965), emerge from the need to study developmental processes that unfold over long periods of time. In psychology, among other fields, some processes such as the development of cognitive skills or brain structure from infancy to adulthood are costly to measure due to the large number of years involved or the expense of measuring certain variables —e.g., through the use of neuroimaging techniques. Such developmental processes are usually studied by collecting repeated measures of variables and analyzing their changes over the years. In an accelerated longitudinal design, participants from different age groups or cohorts are measured over less time and with fewer assessments. For example, instead of measuring participants annually from childhood to early adulthood (e.g., from 5 to 20 years old), in an ALD, different cohorts of participants enter the study at different ages (e.g., at ages from 5 to 18). They are repeatedly measured over two or three years, and then their trajectories are aggregated. This way, the complete developmental period is covered but each age cohort provides only information about a fraction of this period. By combining the overlapping information from each cohort, it is possible to obtain a complete coverage of the target age range in the span of a few years (Duncan et al., 1996; Estrada & Ferrer, 2019). Figure 1 depicts an example of trajectories from a longitudinal design, and the same trajectories under an ALD, where individuals from different cohorts were measured three times with intervals of two years in between.

Figure 1

Trajectories from a standard longitudinal design (left) and an ALD (right)



Note. All cases come from the same population. Subjects in the right panels were measured three times in alternate years. The top and bottom panels represent, respectively, the observed values of two developmental processes that influence each other over the years.

The process of aggregating segments of trajectories is made possible because it relies on the assumption of *cohort equivalence*. We consider two or more cohorts to be equivalent if their trajectories can be characterized by the same set of parameter values, indicating that they come from the same population. This assumption holds true in populations where there are minimal or negligible changes over time (Bell, 1953, 1954; Cáncer et al., 2023; Estrada & Ferrer, 2019).

ALDs have been employed in a variety of fields to investigate longitudinal processes, including recent studies in developmental psychology. For instance, they have been used to examine the behavior of children with autism and the adjustment of their mothers from childhood to pre-adolescence (Benson, 2014), as well as to track changes in brain structure and function over the course of an individual's life (Liu et al., 2021). Other examples are the study of empathy throughout the entire lifespan (Oh at al., 2019) or temporal order memory through childhood (Canada et al., 2020). ALDs have found extensive application in the field of cognitive development, with a particular focus on childhood, adolescence, and early adulthood. Studies by Estrada et al. (2019), Fandakova et al. (2017), Ferrer et al. (2009), Ferrer and McArdle (2004), Green et al. (2017), or Wendelken et al. (2017) have employed these designs to examine cognitive development during the first years of life. It is worth mentioning that these studies are typically based on developmental theories that are multivariate and dynamic, as they involve multiple variables and processes that unfold together over time.

One of the objectives of this manuscript is to help popularize ALDs. Depending on the specific sampling schedule, these designs typically require collecting only 15%-25% of the data points needed in a conventional longitudinal study (i.e., left panels of Figure 1). This substantial reduction implies a great saving of resources. Importantly, in conditions of limited funding, an ALD may be the only viable choice for studying development over a long period of time (e.g., 15 years), as researchers can cover such a period in a study spanning only 3 to 5 years. We aim to contribute to the still limited literature exploring ALDs effectiveness, and to provide specific recommendations on how to apply these designs and analyze the data obtained through them.

Bivariate Latent Change Score Models for Dynamic Processes

Numerous psychological variables are theoretically and empirically interrelated, and they develop over time in an interconnected fashion. A very flexible tool for capturing such interrelated processes is the Bivariate Latent Change Score model (BLCS; Cáncer et al., 2021; Cáncer & Estrada, 2023; Ferrer & McArdle, 2010, Ghisletta & McArdle, 2012; Kievit et al., 2018; McArdle & Hamagami, 2001; McArdle, 2001, 2009). This model, usually specified in a structural equation modeling framework (SEM), has been frequently used in the literature on psychological development and allows examining the interrelations between two variables that unfold over time as a dynamic system, where any given state is defined as a function of its preceding states. This model is a valuable tool for longitudinal studies, as it characterizes the latent change process of a nonlinear system, which allows modelling a wide range of trajectories (Cáncer et al., 2021).

A BLCS model represents the observed state of two variables *X* and *Y* for an individual *i* at any given time *t* as a function of: 1) an initial level (x_{i0} and y_{i0}), 2) the history of latent changes up to that time, which consist of the sum of the *k* previous time intervals (Δx_{ik} and Δy_{ik}), and 3) a measurement error term ($\varepsilon_{i[t]}$). These measurement errors follow a time-invariant normal distribution, with mean zero ($\mu_{\varepsilon x} = \mu_{\varepsilon y} = 0$) and variances ($\sigma^2_{ex}$ and σ^2_{ey}), which are typically allowed to covary (σ_{exey}).

$$X_{i[t]} = x_{i,0} + \sum_{k=1}^{t} \Delta x_{ik} + \varepsilon_{x,i[t]}$$

$$Y_{i[t]} = y_{i,0} + \sum_{k=1}^{t} \Delta y_{ik} + \varepsilon_{y,i[t]}$$

(1)

The history of latent changes is often the focus of these models and, depending on the interests and hypotheses of the researchers, various mathematical specifications are possible for characterizing the mechanism affecting them. For instance, when investigating the cortical and cognitive development of children and adolescents, BLCS models allow to examine how changes in cortical development from one measurement to the next are related to past measurements of both cortical development and cognitive abilities (e.g. Estrada et al., 2019). A very common specification of the latent changes in a bivariate system is[1]:

$$\Delta x_{i[t]} = x_{a,i} + \beta_x \cdot x_{i[t-1]} + \gamma_x \cdot y_{i[t-1]}$$

$$\Delta y_{i[t]} = y_{a,i} + \beta_y \cdot y_{i[t-1]} + \gamma_y \cdot x_{i[t-1]}$$

(2)

Here, the latent change in each variable (Δx and Δy) is a function of: 1) an additive component ($x_{a,i}$ and $y_{a,i}$) that adds a constant amount of change at each new time point, 2) a self-feedback effect that captures the effect from the latent level of the same process at the previous occasion (β) and 3) a coupling effect that captures the effect of the latent level of the other variable at the previous occasion (γ). The initial levels (x_0 and y_0) and the additive components (x_a and y_a) are typically allowed to covariate with the covariance structure defined in Equation 5 (see the next section). The variances of these latent variables capture individual differences within the population.

Figure 2 depicts a path diagram of the BLCS described in Equation 2. Following the previous example, the change in cognitive development at any given time point *t* (e.g. Δx_t) is influenced by 1) a constant amount of change (x_a), 2) an effect of the previous cognitive level (β_x) and 3) an effect of the previous cortical development level (γ_x). In a hypothetical scenario where the initial level of cognitive development positively correlates with the additive component of cortical development ($\sigma_{x0ya} > 0$), individuals with higher cognitive level at t=0 experience a greater linear addition in

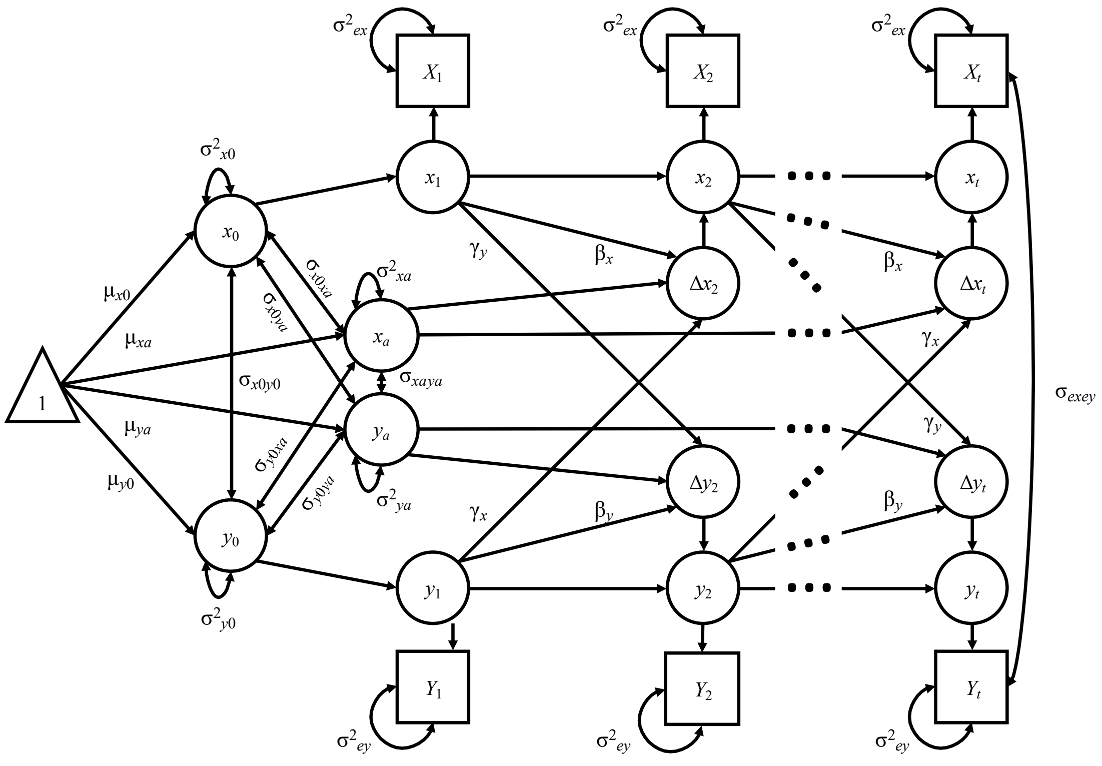
[1] In some cases, dynamic errors are included in the model to consider the influence of random shocks at the latent level, although these stochastic models are not commonly used (see Cáncer et al., 2023 for further explanations of deterministic and stochastic BLCS models)

cortical development at each time point, resulting in higher asymptotic values, that is, higher maximum values that the trajectories reach in the long run (Cáncer et al., 2021).

Some recent examples of longitudinal associations that have been studied using BLCS models include the dynamic relationships between executive functions and anxiety severity (Zainal & Newman, 2021), the interdependence between loneliness and social engagement across the years (Power et al., 2019) or the interplay between emotional regulation and evaluative threat in young adults (Rice et al., 2019).

Figure 2

Path diagram of a Bivariate Latent Change Score model



Continuous Time Dynamic Modeling

The BLCS model described above is defined in discrete time (DT), which implies that time is divided into constant intervals. These models have limited flexibility, as they need the time interval between measurements to remain constant during the entire study, both across participants and across repeated measures for every participant. However, such a regularly spaced time sampling is rarely found in empirical studies. This is why longitudinal research in psychology can benefit greatly from the use of continuous-time (CT) modeling, which has been proposed in recent years as a robust and effective approach to longitudinal data analysis (de Haan-Rietdijk et al., 2017; Deboeck & Preacher, 2016; Oud & Jansen, 2000, Voelkle et al., 2012; Voelkle & Oud, 2015).

In contrast to the DT framework, a CT model considers time as a continuous variable and defines the changes in a system for an infinitesimal time interval (*dt*, instead of Δt, Ryan et al., 2018; Voelkle et al., 2012). As a result, CT models are considered to account for the process between observations, whereas DT models are assumed to reflect only changes occurring at discrete time points (Oud & Delsing, 2010). A highly practical advantage is that CT models allow researchers to collect data at varying intervals, which can increase efficiency and provide more precise estimates of change (Ryan et al., 2018). In contrast, DT models require consistent time intervals between measurements, which can limit the sampling design options. Another important advantage of using CT models is that their parameters are independent of the specific time intervals used in any given study, and therefore they make it possible to compare parameters obtained in studies that use different time intervals (Voelkle, 2012).

In CT modeling, differential equations are used to describe the change of the variables of interest. The latent changes specified in DT in Equation 2 can be expressed in a CT framework as a first order ordinary differential equation containing the same dynamic parameters (although with a somewhat different interpretation, since effects are defined for an infinitesimally small time lag):

$$\frac{dx_i(t)}{dt} = x_{a,i} + \beta_x \cdot x_i(t) + \gamma_x \cdot y_i(t)$$

$$\frac{dy_i(t)}{dt} = y_{a,i} + \beta_y \cdot y_i(t) + \gamma_y \cdot x_i(t) \tag{3}$$

These CT values provide information for computing the analogue DT values at any desired time interval (Δt). For more detailed information of the relation between CT and DT parameters and their interpretation, see Voelkle and Oud (2015) and Voelkle et al. (2012).

Because in ALDs different individuals are expected to be measured at different ages, and (as in any longitudinal study) the measurement occasions are almost never evenly spaced (i.e., time intervals are not constant), CT models appear to be a very suitable tool for modeling this data. In fact, previous works have shown that CT models perform better in recovering generating parameters of a univariate system, particularly when observations are unevenly spaced, compared to LCS models in DT (Estrada & Ferrer, 2019).

State-Space Models

In recent years, various modeling frameworks have been proposed in the psychological literature for estimating CT models including latent variables. A very interesting one is State-Space Modeling (SSM), which was specifically designed for dynamic systems with longitudinal data (Chow et al., 2010; Hunter, 2018; Voelkle & Oud, 2015). In this study, we use a state-space model in continuous time to analyze the recovery of trajectories of bivariate systems in ALDs. SSMs consist of two equations: the *state equation* and the *output equation*. For a deterministic CT-BLCS, the *state* (or *transition*) *equation* describes change in a vector of the four latent variables for an infinitesimally brief time interval (*dt*):

$$\frac{d}{dt}\begin{bmatrix} x_{l,i} \\ y_{l,i} \\ x_{a,i} \\ y_{a,i} \end{bmatrix} = \begin{bmatrix} \beta_x & \gamma_x & 1 & 0 \\ \beta_y & \gamma_y & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{l,i} \\ y_{l,i} \\ x_{a,i} \\ y_{a,i} \end{bmatrix} \tag{4}$$

where x_l and y_l represent the time-varying latent level of each variable for each individual, and x_a and y_a represent the time-invariant linear components—related to the additive component in the SEM BLCS model described previously—, which influence the latent levels of their corresponding process (either *x* or *y*). Here, latent states at t=0 (x_{i0} and y_{i0}), are defined by 1) a latent mean vector within the means from the additive components (x_a and y_a) and 2) a latent covariance matrix that informs about the interrelations between the four latent variables which follows a multivariate normal distribution (see Cáncer & Estrada, 2023):

$$\begin{bmatrix} x_0 \\ y_0 \\ x_a \\ y_a \end{bmatrix} \sim N\left(\mu = \begin{bmatrix} \mu_{x0} \\ \mu_{y0} \\ \mu_a \\ \mu_a \end{bmatrix}, \Sigma = \begin{bmatrix} {\sigma_{x0}}^2 & & & \\ \sigma_{x0y0} & {\sigma_{y0}}^2 & & \\ \sigma_{x0xa} & \sigma_{y0xa} & {\sigma_{xa}}^2 & \\ \sigma_{x0ya} & \sigma_{y0ya} & \sigma_{xaya} & {\sigma_{ya}}^2 \end{bmatrix} \right) \tag{5}$$

The *output equation* is similar to the measurement model from structural equation models (SEM; Chow et al., 2010; Hunter, 2018). It connects the latent variables defined in continuous time (Equation 4) with the observed measurements (*X* and *Y*) for each subject *i*:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_{l,i} \\ y_{l,i} \\ x_{a,i} \\ y_{a,i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,i} \\ \varepsilon_{y,i} \end{bmatrix} \tag{6}$$

Therefore, the SSM consist of a *latent, unobservable equation* that captures the time-lagged dynamics of both processes and a *measurement equation* linking the unobserved and observed variables, which includes measurement errors (ε). These errors have zero mean and a covariance matrix with three parameters (σ^2_{ex}, σ^2_{ey} and σ_{exey}). For a detailed explanation of the mathematical specification of a complete general SSM, see Chow et al. (2010) and Hunter (2018).

Purpose of the study

Previous research has shown that the dynamics underlying ***univariate*** developmental processes can be adequately recovered in accelerated longitudinal designs, both when the different cohorts are equivalent (Estrada & Ferrer, 2019; Estrada et al., 2020), and when they are not (Cáncer et al., 2023; Estrada et al., 2021; Miyazaki & Raudenbush, 2000). However, as stated previously, psychological phenomena are typically conceived as multivariate, and the dynamics of two or more processes that unfold over time are considerably more difficult to recover.

Therefore, the present study aims to examine the extent to which the features of a ***bivariate*** developmental system (e.g., co-development of reading and mathematics abilities, or cognitive ability and cortical thickness) can be recovered when data are gathered through accelerated longitudinal designs. To the best of our knowledge, this is the first exploration of the performance of ALDs with bivariate systems. For this reason, we started in a relatively simple scenario and assumed the different cohorts to be equivalent. This implies that they come from the same population and share a common developmental trajectory.

Particularly, we explore the effectiveness of the bivariate latent change score model described above, estimated in continuous time as a state-space model. We evaluate the performance of this model in various ALDs, with the aim of identifying the optimal sample designs and sizes for recovering the true trajectories while remaining cost-effective. Hopefully, the results from this study will provide novel insights and guidelines for designing future substantive studies.

Method

Monte Carlo Study

We generated repeated measurements of two interrelated latent processes that unfold over time for 15 years, according to the SSM defined in Equations 4 and 5. We selected parameters for generating the populational data describing trajectories that are typical of the development of cognitive abilities from childhood to early adulthood. Such trajectories typically follow exponential decelerated growth, as they show rapid increases at the beginning of the time period considered, and progressively decelerate until they reach a peak during early adulthood (McArdle et al., 2002).

Our specific generating parameters were adapted from a previous paper that examined the effectiveness of the BLCS model for developmental processes (Cáncer & Estrada, 2023). In this previous work, a set of generating parameters was selected by conducting a literature review on developmental processes and identifying typical values for the parameters in a bivariate LCS system. However, in that study, the model was specified in the classic DT SEM version (i.e., with constant time intervals, $\Delta t = 1$). In order to generate a bivariate system unfolding in truly continuous time, we rescaled those parameters to plausible values in CT. For this, we simulated scores for multiple samples with constant time intervals between measurements according to the original DT parameters. To facilitate the estimation of the model and subsequent interpretations, we standardized the resulting scores with respect to the first measurement occasion: for all the scores, we subtracted the mean of the first measurement occasion, and then divided by the standard deviation of such first occasion. Later, we fitted a continuous time SSM to the resulting datasets and saved the corresponding CT values. We repeated this process 20 times (i.e., 20 samples of size 250) and averaged the estimates from the SSMs. These averaged estimates were then used for the Monte Carlo simulation in our study. They are reported in Table 1. The graphs from Figure 1 depict examples of the generated trajectories in the two variables from 5 to 20 years of age in one of the simulated samples (n = 100).

Table 1.

Generating population parameters values for the simulation study

Parameter	Value in DT ($\Delta t = 1$)	Value in CT
Initial level means		
μ_{x0}	-0.8	-0.022
μ_{y0}	-0.2	0.018
Initial level variances		
σ^2_{x0}	0.3	1.014
σ^2_{y0}	0.6	1.037
Additive component means		
μ_{xa}	0.8	2.279
μ_{ya}	1	1.105
Additive component variances		
σ^2_{xa}	0.02	0.095
σ^2_{ya}	0.08	0.174
Self-feedbacks		
β_x	-0.35	-0.461
β_y	-.25	-0.313
Couplings		
γ_x	0.1	0.212
γ_y	0.2	0.211
Covariances between initial levels and additive components		
σ_{x0y0}	0.212	0.513
σ_{xaya}	0.016	0.051
σ_{x0xa}	0.046	0.186
σ_{x0ya}	0.031	0.084
σ_{y0xa}	0.033	0.094
σ_{y0ya}	0.131	0.254
Correlations between initial levels and additive components		
r_{x0y0}	0.5	0.5
r_{xaya}	0.4	0.4
r_{x0xa}	0.6	0.6
r_{x0ya}	0.2	0.2
r_{y0xa}	0.3	0.3
r_{y0ya}	0.6	0.6
Variances, covariance and correlation between measurement errors*		
σ^2_x	0.1	0.333
σ^2_y	0.2	0.335
σ_{xy}	0.02	0.0473
r_{xy}	0.14	0.14

Note: Measurement error values changed to maintain the rate of error once the data was standardized. Values in DT were based on Cáncer and Estrada (2023).

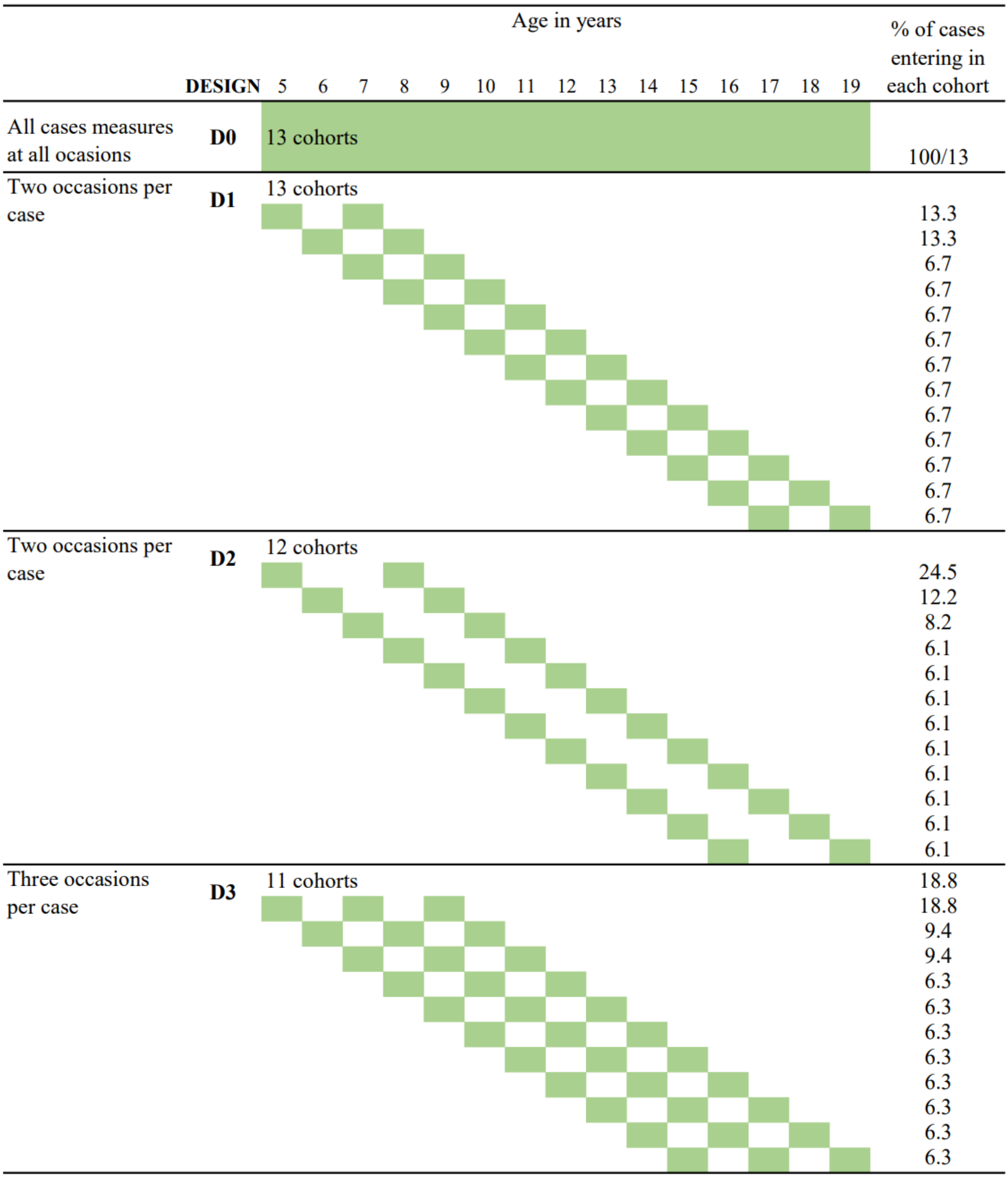
Regarding the simulation conditions of the study, we manipulated two factors: (1) the sampling schedule and (2) the sample size. For the first one, we sampled the generated trajectories of both variables according to three ALDs. Namely, we chose three sampling designs that could be applied in less than five years. Each design implied a different duration of the study, average time interval between assessments, and number of measurement occasions per person. Figure 3 depicts the sampling schemes evaluated.

Design 0 represents a non-cohort design, where all participants are measured every year. This is not an ALD, but a traditional longitudinal study spanning 15 years. We included this condition as a benchmark for comparing results from the remaining ALDs. Design 1 includes two measurements per case, sampled two years apart. Therefore, the sample collection would last three years. In Design 2, the two measurements are separated three years apart, so the study would take four years. Finally, in Design 3, all cases are measured three times, sampled two years apart, so it spans five years for each cohort. The examination of these designs is of interest, as Designs 2 and 3 incorporate a fewer number of cohorts, while Designs 1 and 2 have a smaller number of assessments and therefore shorter time for sample collection.

Because these developmental processes are better estimated when the data density is as high as possible at the region with the greatest curvature (Mistler & Enders, 2012; Rhemtulla & Hancock, 2016), we assigned more subjects to cohorts covering the age range of 5 to 9 years. We sampled the data points for each subject randomly uniformly at any given week within the corresponding year (as depicted in Figure 1). That is, in the years scheduled for measurement, the individuals were measured once a year in a randomly selected week, with all weeks of the year having the same probability.

Figure 3

Sampling schedules for the accelerated longitudinal design (ALD) examined



The second factor that we manipulated was the sample size. We included seven sample conditions: 100, 175, 250, 325, 400, 475 and 550. Therefore, the study included 7 x 4 = 28 simulation conditions. In each of them, we generated 200 replications (i.e., 28 x 200 replications in total).

Data Analysis

In each of the simulated samples, we estimated a state space model in continuous time (SSM-CT). We used the *OpenMx* package in *R* (Boker et al., 2018; Hunter, 2018; Neale et al., 2016) and its functions *mxExpectationStateSpaceContinuousTime* and *mxFitFunctionML*. These functions utilize a set of recursive algorithms termed a "hybrid" Kalman Filter to estimate each subsequent state vector and state covariance matrix. This is done through a prediction step followed by an update step from the observed measurements. The Kalman Filter then adjusts the parameters of the model using Maximum Likelihood prediction error decomposition to reduce the prediction error (Hunter, 2018; Kalman, 1960; Neale et al., 2016; You et al., 2020). An R script for specifying and fitting the model to any sample can be found on this OSF repository: https://osf.io/7jzsv/?view_only=5ccfe6cfd2d144969572f37677f4874f. The computation time required to fit a model to a sample depends largely on the sample size and the number of measurements available for each case. In our study, the condition with the most data points included 550 cases with 15 measurements each (Design 0). On a commercial computer with regular computing power (i.e., a laptop), the model took no more than 5 minutes to converge for any of the samples.

Results

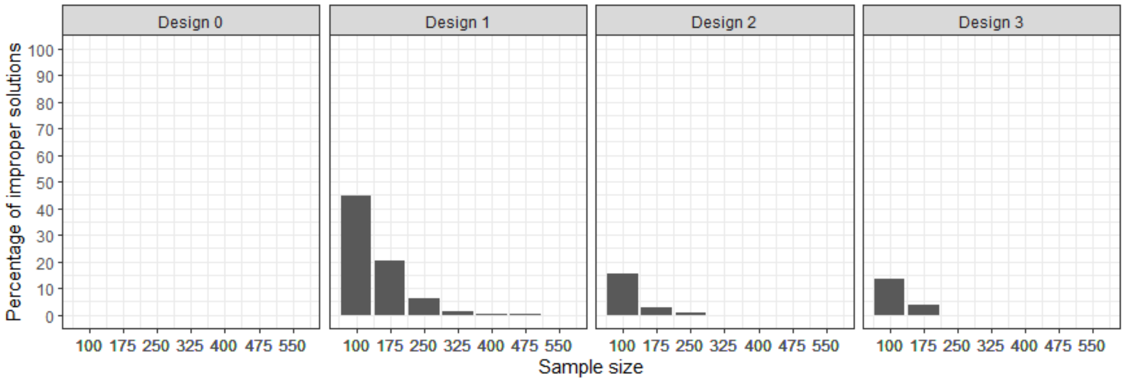
Due to the large amount of information generated in the simulation, this paper presents the results in graphical format. The tables with all the numerical results, as the R code for estimating CT-BLCS models are available at: https://osf.io/7jzsv/?view_only=5ccfe6cfd2d144969572f37677f4874f.

Improper solutions

During the estimation of the model, convergence problems were found in some replications, leading to improper solutions. We defined improper solution as outcomes where the estimated parameters contained missing values (NAs) and the standard errors exhibited extreme values. The number of replicates that resulted in an improper solution varied across different conditions. Figure 4 depicts the rates of improper solutions.

Figure 4

Percentage of improper solutions across conditions



According to Figure 4, the majority of improper solutions occurred in conditions with a sample size of fewer than 250 subjects in all ALDs. Design 0, which presents a conventional 15-year-longitudinal design, did not lead to any improper solution. In contrast, Design 1, which involved measuring each subject twice in alternate years, was more likely to present converge problems. For example, 44.5% of improper solutions were found in samples of 100 participants. Only the samples with proper solutions in each condition were considered for subsequent analyses.

Estimation Accuracy, Variability and Coverage

To evaluate the model performance, we calculated three indicators: relative bias (*RB*), standard deviation of relative bias (*SDRB*) and Coverage.

To evaluate the accuracy of the model, we calculated the relative bias (*RB*) for each parameter using the formula $RB = (\bar{\theta}_{est} - \theta)/|\theta|$, where θ represents the true parameter value and $\bar{\theta}_{est}$ represents the average estimated value of the parameter from all replications in a given condition. A RB value closer to zero indicates unbiased estimates, positive values suggest overestimation, and negative values imply underestimation. Previous research has considered estimates to be substantially biased when |*RB*| > 0.10 (Flora & Curran, 2004; Rhemtulla et al., 2012). In cases where the true value of an estimate is very close to zero (i.e., initial levels means μ_{x0} and μ_{y0}), we report the raw bias $B = (\bar{\theta}_{est} - \theta)$, instead of the *RB* indices. This was done to prevent their *RB* indices from being distorted by dividing by such a small value.

To assess the variability of the parameter estimates, we calculated the standard deviation of relative bias (*SDRB*) across all the samples in each condition by computing $SDRB = SD[(\theta_{est} - \theta)/\theta]$. This index allowed expressing the estimation inefficiency in the same scale for all parameters. *SDRB* values are always positive, and lower values indicate less variability in the estimated values within a given condition, indicating greater efficiency.

Finally, we calculated the Coverage index, which represents the proportion of 95% confidence intervals around the estimated parameter value that include the true parameter value. An optimal Coverage is achieved when at least 95% of the intervals contain the true parameter within their boundaries, while an adequate Coverage is obtained when at least 90% of the intervals include the true parameter (Collins et al., 2001; Enders & Peugh, 2004).

Given that the BLCS models consist of 21 parameters, we present the indices by grouping them into distinct parameter sets. Initially, we examine the recovery of the dynamic parameters (β and γ). Subsequently, we present the indices related to the means and variances of the latent variables: the initial levels (x_0 and y_0) and the additive components (x_a and y_a). Next, we analyze the covariances and correlations among these variables. Lastly, we analyze the measurement errors variances and covariance recoveries.

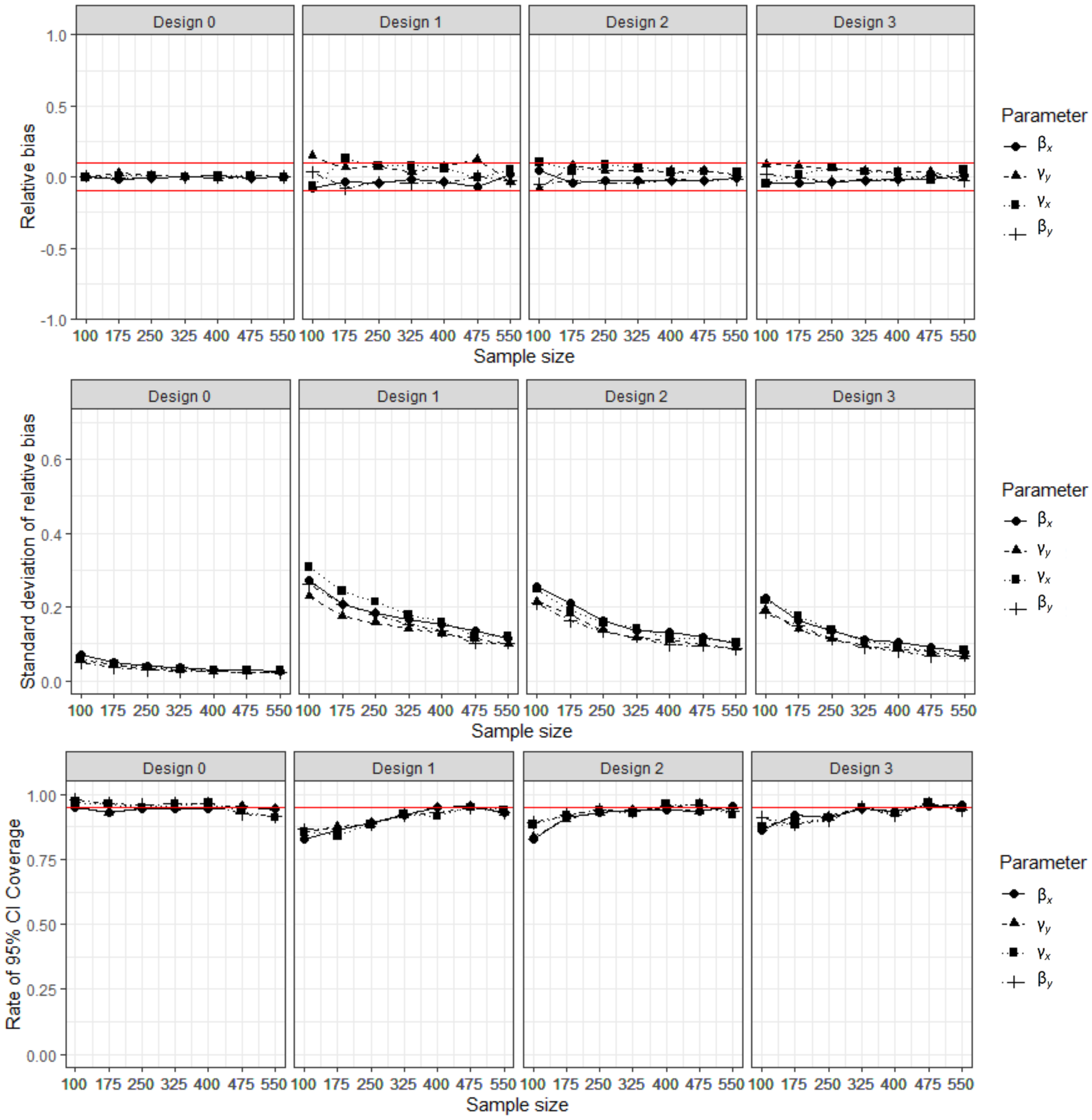
Dynamic parameters

As depicted in Figure 5, all parameters were adequately recovered across all ALDs when the sample sizes exceeded 250 subjects. A notable reduction of bias was observed as the sample size increased, which was expected since the model has more information for the estimation. For sample sizes smaller than 250, Design 1 showed a poor performance. The self-feedback parameters (β_x and β_y) were underestimated, whereas the couplings (γ_x and γ_y) were overestimated in those cases. In samples of size 100, none of the ALDs produced satisfactory results. Notably, Design 2, as depicted in the third column of Figure 5, exhibited superior performance compared to Design 1 and yielded comparable results to Design 3. This is an interesting finding because Design 2 includes one less assessment and one less year of data collection compared to Design 3.

Regarding the variability of the estimates, a similar pattern was observed across all three ALDs, where the variability decreased with larger sample sizes. Concurrently, the coverage levels demonstrated a similar trend in all three designs, attaining adequate (>.90) and optimal (>.95) levels with increasing sample size. Hence, confidence intervals are more likely to enclose the true parameter as the sample is larger.

Figure 5

Relative bias, Standard deviation of relative bias and Coverage indices for dynamic parameters across all conditions



Note: The red lines depict |RB|=.10. for the relative bias panels and 95% coverage for the rate of 95% CI coverage panels.

Latent variables: initial levels and additive components

Similar to the dynamic parameters, the estimation of the means and variances of the initial levels and additive components were adequately recovered in Designs 2 and 3, particularly in larger sample sizes (n > 250), as depicted in Figure 6. With smaller sample sizes, Designs 1 and 2 overestimated the variances of the additive components (σ_{xa}^2 and σ_{ya}^2).

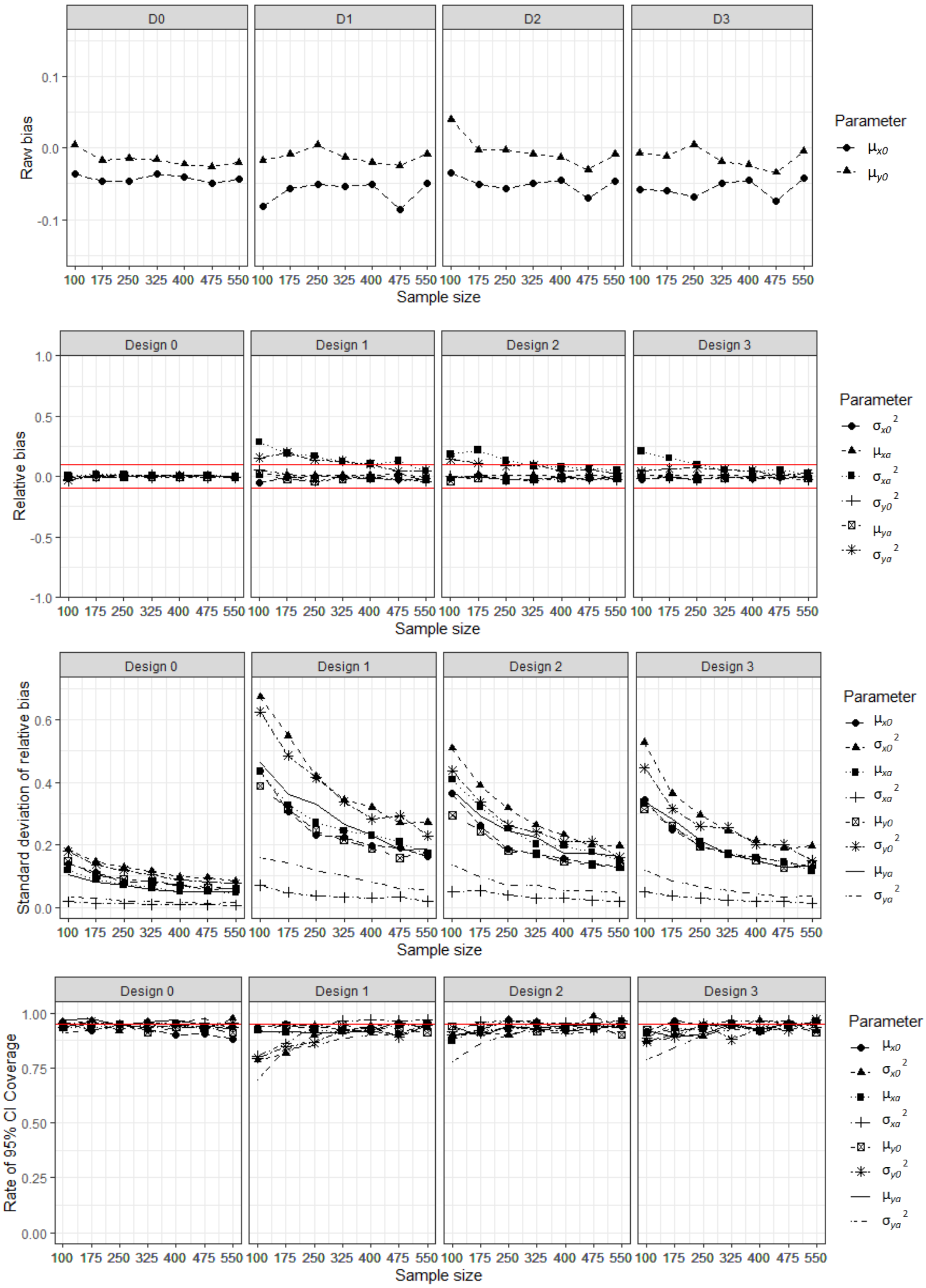
Because the means of the initial levels had a true value very close to zero, we computed the raw bias, instead of the relative bias, for them. Consequently, there is no specific criterion to assess their estimation accuracy. Nonetheless, a comparison was be made with Design 0, which had demonstrated favorable outcomes. Thereby, the discrepancy between the estimated and true values did not appear to be substantial in any condition.

This set of parameters showed higher SDRB values, which implies higher estimation inefficiency. This is particularly evident in Design 1, where the variances of the initial levels (σ_{x0}^2 and σ_{y0}^2) show higher levels of inefficiency. It is also noteworthy that, in terms of coverage, none of the variances exhibited satisfactory results with n = 100, and with n < 250 in Design 1.

Coverage exhibited a similar pattern to the dynamic parameters, ranging between values of .90 and .95. Nevertheless, it is worth noting that the variances of initial levels (σ_{x0}^2 and σ_{y0}^2) and additive components (σ_{xa}^2 and σ_{ya}^2) displayed lower levels in the small sample sizes of Design 1.

Figure 6

Raw bias, relative bias, Standard deviation of relative bias and Coverage indices for mean and variances of latent variables across all conditions

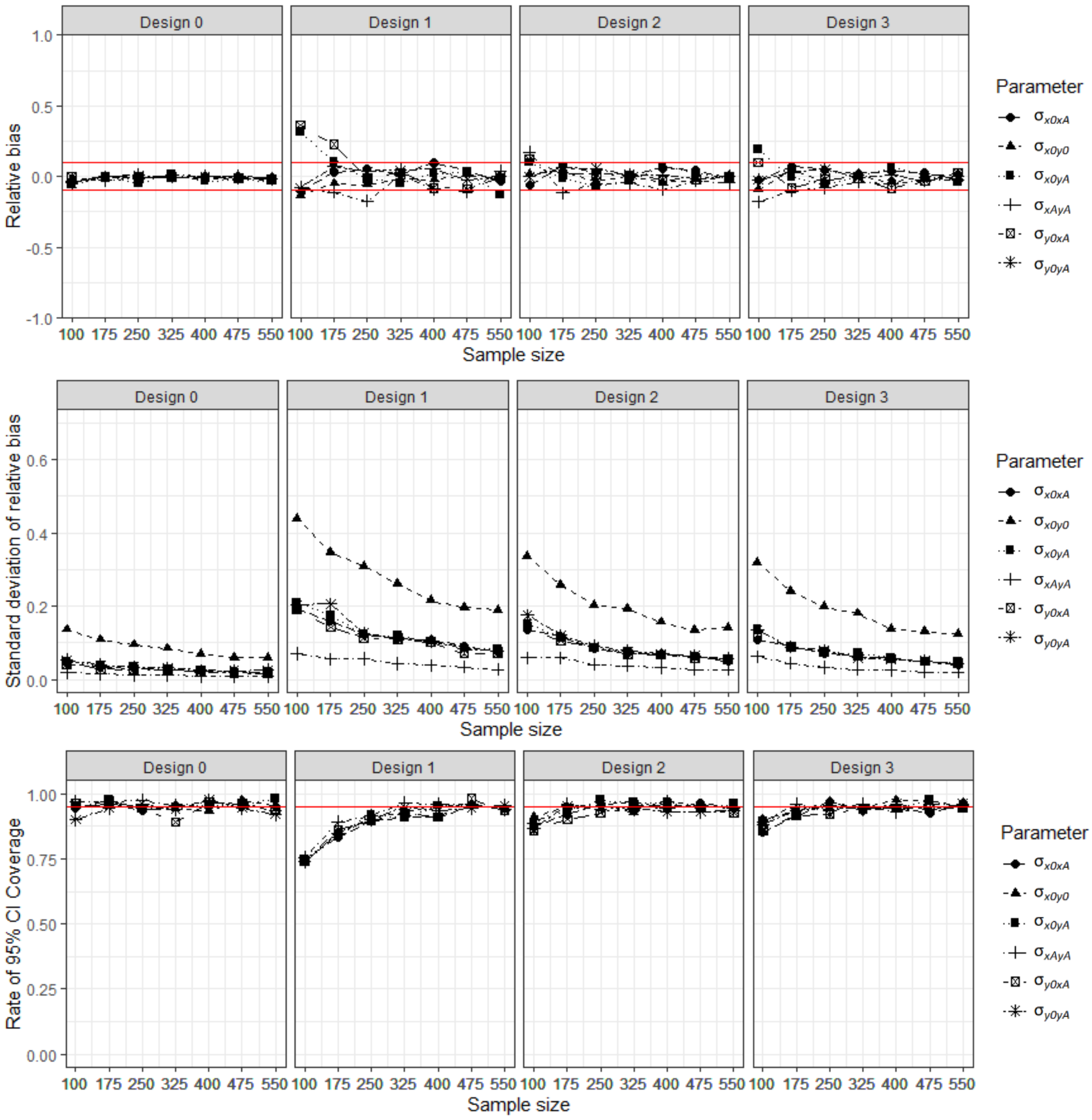


Note: The red lines depict |RB|=.10. for the relative bias panels and 95% coverage for the rate of 95% CI coverage panels.

Regarding the covariances between latent variables, Figure 7 depicts the recovery of covariances between initial levels and additive components. Once again, Design 1 presented a poor performance with $n < 250$ regarding relative bias and coverage values. Additionally, a sample size of only 100 participants appeared to be inadequate, regardless of the ALD utilized.

Figure 7

Relative bias, Standard deviation of relative bias and Coverage indices for covariances between latent variables across all conditions



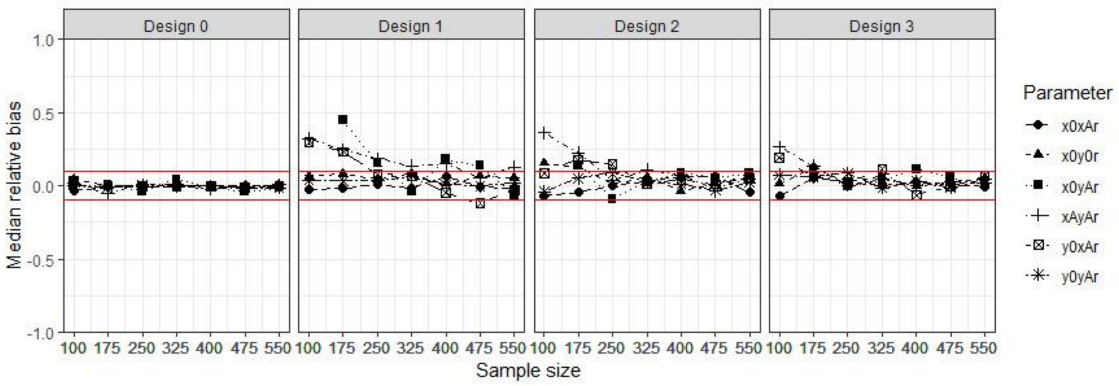
Note: The red lines depict |RB|=.10. for the relative bias panels and 95% coverage for the rate of 95% CI coverage panels.

Additionally, the estimation inefficiency (i.e., *SDRB*) for the covariance between the initial levels (σ_{x0y0} = 0.51) is distinctively higher compared to the other parameters in all conditions. This observation may be attributed to the relatively higher value of the covariance itself, which stands out among the other covariances (See Table 1). It is important to note that the magnitude of covariances is heavily influenced by the metric of variances. Thus, comparing the values of covariances alone does not provide a definitive interpretation. To gain a more accurate understanding, we also present correlations, which can be directly compared as they are expressed on the same scale.

Initially, the relative bias for correlations obtained displayed abnormal values at smaller sample sizes (see Appendix for further details). Consequently, we computed the median (instead of mean) relative bias (depicted in Figure 8), $MRB = (Median(\theta_{est}) - \theta)/\theta$. As expected, the correlation estimates were more accurate in larger samples. In this case, Design 3 showed better estimation accuracy than Designs 1 and 2 for smaller samples sizes (n<175)

Figure 8

Median relative bias for correlations across all conditions



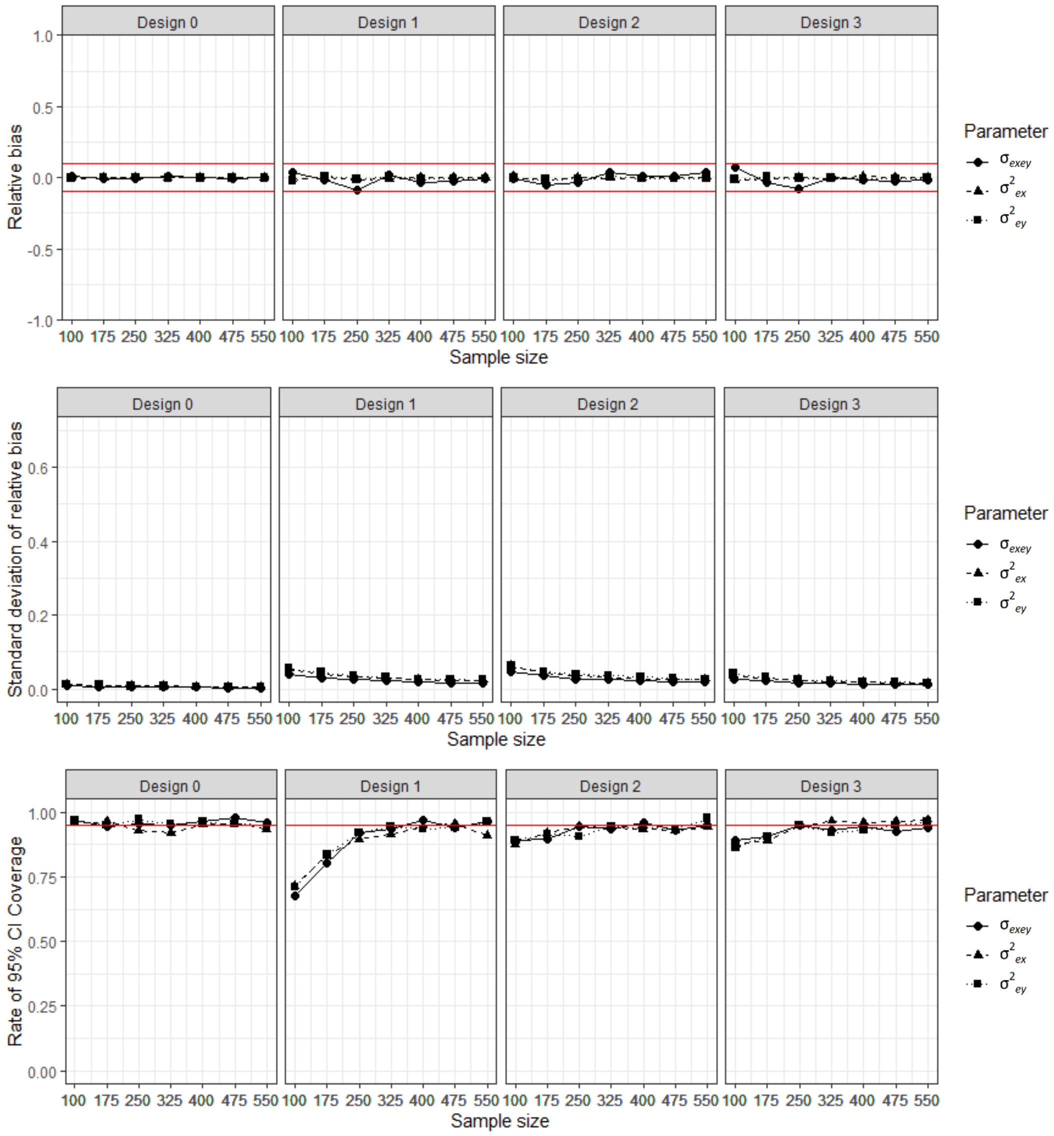
Notes: The red lines depict |MRB|=.10. for the Median relative bias.

Measurement error

The recovery of the measurement error was very good across all conditions, regardless of the ALD. The results are depicted in Figure 9. We found very small variability across replications (SDRB panel of Figure 9). However, the coverage of the measurement error did not achieve adequate values in samples below 250 cases, with Design 1 demonstrating the poorest performance in this regard. This is unexpected, given that coverage is not usually bad when estimates are unbiased. Such an outcome could be attributed to an underestimation of standard errors, leading to excessively narrow confidence intervals. Consequently, these intervals may exclude the true parameter value more frequently than expected, even when the point estimates are highly accurate in the replications.

Figure 9

Relative bias, Standard deviation of relative bias and Coverage indices for measurement error across all conditions



Note: The red lines depict |RB|=.10. for the relative bias panels and 95% coverage for the rate of 95% CI coverage panels.

Discussion

Summary of Findings

In this study, our aim was to evaluate the use of a bivariate latent change score model estimated in continuous time (CT-BLCS) specified in a state-space modeling framework (SSM) for recovering the trajectories of two interdependent developmental processes measured through different accelerated longitudinal designs (ALDs), and to provide specific recommendations for the design of future studies. To achieve this, we conducted a simulation study with three plausible conditions of sampling design and seven conditions of sample size, ranging from 100 to 550.

Several noteworthy observations have emerged from our analysis. First, consistent with expectations, both accuracy and efficiency (i.e., lower variability across replications in a condition) of the parameter estimates improved as the sample size increased, for all ALDs. However, we found a clear difference between the use of smaller (100 and 175) and larger (≥250) sample sizes. This was particularly evident for the Design 1, which exhibited inadequate performance with samples of 100 and 175 cases. Furthermore, most of the improper solutions came from these conditions, which indicates that the model had more difficulties in estimating the parameters with such limited data.

Second, Designs 2 and 3 yielded highly satisfactory results, particularly with samples larger than 175. Finding a similar performance between these two designs is very interesting because Design 2 includes one fewer assessment than Design 3 and it costs only three years to complete, whereas Design 3 needs five years. In other words, Design 2 leads a similar estimation of true parameters with less information than Design 3, making it the most cost-effective design.

Lastly, the different parameters followed a similar pattern of recovery across all conditions. However, it is noteworthy that the variances of the initial components show a slightly lower estimation efficiency (i.e., higher SDRB) compared to the rest of parameters. Additionally, the dynamic parameters and the error variances and covariance showed minimal variation between the different conditions, demonstrating the highest accuracy and efficiency in their recovery.

Theoretical and Methodological considerations

To the best of our knowledge, this study is the first investigation about the performance of accelerated longitudinal designs for the study of bivariate systems. By focusing on the bivariate context, we aimed to fill this gap in the existing literature and shed light on the potential advantages and challenges associated with employing ALDs in studying such systems. Investigating the performance of accelerated longitudinal designs for bivariate systems is very important because these designs offer a notable advantage in terms of cost-effectiveness compared to conventional longitudinal studies. The utilization of ALDs enables the feasibility of studying research questions that would otherwise be financially unattainable.

We note that these results were found in empirically plausible scenarios. Our approach involved sampling assessments for each subject randomly throughout the year, simulating real-world scenarios where measurements are not obtained at precisely constant intervals, resulting in varying time intervals between assessments. This has significant implications in the successful recover of trajectories, as prior studies have consistently shown large biases in the estimation of all parameters of BLCS when measurement intervals differ across cases in discrete time (O'Rourke et al., 2021). Furthermore, these authors identified convergence problems associated with these DT models under varying time metric.

One notable aspect in our study was the occurrence of improper solutions, particularly in samples with fewer than 175 participants and in Design 1, which involves two measurements spaced one year apart. A plausible explanation for this observation is the challenge faced by the model in fitting trajectories with limited data. Having two measurements of a developmental trajectory that are very close (less than one year apart) may not provide sufficient information for accurate estimation. In contrast, Design 2, also featuring two measurements but separated by two years, allows for more significant changes in the latent trajectory, providing more informative data for estimation.

Our findings align with previous simulation studies, providing further evidence that utilizing continuous time models can mitigate biases in parameter estimation when dealing with unequal measurement intervals (Estrada & Ferrer, 2019; Voelkle & Oud, 2013). Therefore, we demonstrate that adopting a continuous-time approach offers researchers the flexibility to accommodate exact measurement times of nonlinear processes, as it ensures accurate parameter estimation within the Latent Change Score model framework.

Limitations and future directions

In this study, our focus was on examining a scenario characterized by cohort equivalence, where trajectories from different cohorts can be described using the same set of parameter values. However, cohorts may differ in some aspects of the developmental trajectories due to a number of reasons. This is because, although in ALDs such as the ones examined here cohorts are separated only a few years apart, some of them could be affected by various environmental factors affecting their development (e.g., changes in the educational laws in the country, an international pandemic, or gradual changes in ethnic composition of the population, among others.). In this paper, we have examined a relatively simple scenario in which cohorts were equivalent, but it is important for subsequent investigations to explore strategies for handling different types of non-equivalent cohorts. Previous works have both analyzed the effect of such non-equivalence in univariate systems, and proposed tools for addressing these situations (Cáncer at al., 2023; Estrada et al., 2021; Estrada & Ferrer, 2019; Miyazaki & Raudenbush, 2000). Understanding how to address such non-equivalence in the context of bivariate developmental processes can provide valuable insights for refining and expanding the application of longitudinal models in new contexts.

In any longitudinal study, some degree of participant attrition is expected, leading to some proportion of unplanned missing data. It is unknown how this additional unplanned missing data affects the estimates of statistical models fitted to data obtained in a scenario with an already high proportion of *planned* missing data, such as those found in ALDs. Future research should explore this important issue.

Another challenge for future studies would be the application of ALDs in other bivariate dynamic processes with other non-exponential trajectories or spanning other life periods, such as old age (e.g. Lee et al., 2023). In these cases, we might encounter not only differences between cohorts, but also large individual variability related to factors such as medical conditions or life events that could lead to different forms of the trajectories.

While the present study focuses on bivariate developmental systems, there is potential for extending this approach to multivariate systems. However, it is important to acknowledge that as the number of processes included increases, so does the complexity of the model. For instance, a univariate LCS model typically includes 7 parameters, while the corresponding bivariate model includes 21 parameters. Introducing a third or fourth longitudinal variable further can increase the number of parameters to 42 and 70, respectively. To the best of our knowledge, such models are very rarely applied in the literature, either in discrete or continuous time. This is probably due to their considerable mathematical complexity, the need for very large samples, and the very likely convergence problems. Although specifying these models is, in principle, feasible, their application must be accompanied by a theoretical basis that justifies the inclusion of additional elements and the associated complexities.

Conclusions and recommendations

Studying interdependent developmental processes can be challenging for researchers, particularly when they unfold over very long periods of time. The present study is the first exploration of the efficacy of a bivariate latent change score model in continuous time in the framework of accelerated longitudinal designs. Considering the findings from our Monte Carlo study, we offer the following recommendations for the design of future empirical studies applying ALDs for developmental variables:

- Our results clearly support the appropriateness of accelerated longitudinal designs for examining the dynamics of two interrelated developmental variables. Therefore, we recommend using these designs, as they lead to a substantial decrease in the time and financial costs required for conducting developmental research.
- Taking two assessments separated two years apart (i.e., a study of four years in total; Design 2) leads to very good estimation of true parameters with samples of 250 cases or more.
- If it is not possible to wait for two years before taking the second measurement, and the researcher needs to take two measurements separated two years apart (i.e., Design 1, spanning three years in total instead of four), a sample of at least 325 participants should be used.

- In the case of having not having access to very large samples, it is more appropriate to plan a design including two measurements separated one year apart (i.e., Design 3, spanning five years in total). Nonetheless, it is preferable to avoid a sample size smaller than 175.

By adhering to these recommendations, researchers can optimize their ALDs and enhance the accuracy and efficiency of their investigations. We hope this work contributes to the popularization of accelerated longitudinal designs and continuous time modeling in developmental psychology.

References

Bell, R. Q. (1953). Convergence: An Accelerated Longitudinal Approach. *Child Development*, *24*(2), 145–152. https://doi.org/10.2307/1126345

Bell, R. Q. (1954). An Experimental Test of the Accelerated Longitudinal Approach. *Child Development*, *25*(4), 281–286. https://doi.org/10.2307/1126058

Benson, P. R. (2014). Coping and Psychological Adjustment Among Mothers of Children with ASD: An Accelerated Longitudinal Study. *Journal of Autism and Developmental Disorders*, *44*(8), 1793–1807. https://doi.org/10.1007/s10803-014-2079-9

Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R., Estabrook, R., Bates, T. C., Mehta, P. D., von Oertzen, T., Gore, R. J., Hunter, M. D., & Hackett, D. C. (2018). *OpenMx User Guide*. https://openmx.ssri.psu.edu/documentation

Canada, K. L., Pathman, T., & Riggins, T. (2020). Longitudinal development of memory for temporal order in early to middle childhood. *The Journal of genetic psychology*, *181*(4), 237-254. https://doi.org/10.1080/00221325.2020.1741504

Cáncer, P. F., & Estrada, E. (2023). Effectiveness of the Deterministic and Stochastic Bivariate Latent Change Score Models for Longitudinal Research. *Structural Equation Modeling: A Multidisciplinary Journal*, *0*(0), 1–15. https://doi.org/10.1080/10705511.2022.2161906

Cáncer, P. F., Estrada, E., & Ferrer, E. (2023). A Dynamic Approach to Control for Cohort Differences in Maturation Speed Using Accelerated Longitudinal Designs. *Structural Equation Modeling: A Multidisciplinary Journal*, *0*(0), 1–17. https://doi.org/10.1080/10705511.2022.2163647

Cáncer, P. F., Estrada, E., Ollero, M. J. F., & Ferrer, E. (2021). Dynamical Properties and Conceptual Interpretation of Latent Change Score Models. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.696419

Chow, S.M., Ho, M. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and Differences Between Structural Equation Modeling and State-Space Modeling Techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(2), 303–332. https://doi.org/10.1080/10705511003661553

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6,* 330–351. https://doi.org/10.1037/1082-989X.6.4.330

de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L., & Hamaker, E. L. (2017). Discrete- vs. Continuous-Time Modeling of Unequally Spaced Experience Sampling Method Data. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01849

Deboeck, P. R., & Preacher, K. J. (2016). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 61-75. https://doi/abs/10.1080/10705511.2014.973960

Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological Methods, 1*(3), 236–248. https://doi.org/10.1037/1082-989X.1.3.236

Enders, C. K., & Peugh, J. L. (2004). Using an EM covariance matrix to estimate structural equation models with missing data: Choosing an adjusted sample size to improve the accuracy of inferences. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 1–19. https://doi.org/10.1207/S15328007SEM1101_1

Estrada, E., Bunge, S. A., & Ferrer, E. (2021). Controlling for cohort effects in accelerated longitudinal designs using continuous- and discrete-time dynamic models. *Psychological Methods*, *28*¸ 359-378. https://doi.org/10.1037/met0000427

Estrada, E., & Ferrer, E. (2019). Studying developmental processes in accelerated cohort-sequential designs with discrete- and continuous-time latent change score models. *Psychological Methods*, *24*, 708–734. https://doi.org/10.1037/met0000215

Estrada, E., Ferrer, E., Román, F. J., Karama, S., & Colom, R. (2019). Time-lagged associations between cognitive and cortical development from childhood to early adulthood. *Developmental Psychology*, *55*, 1338–1352. https://doi.org/10.1037/dev0000716

Estrada, E., Hamagami, F., & Ferrer, E. (2020). Estimating Age-Based Developmental Trajectories Using Latent Change Score Models Based on Measurement Occasion. *Multivariate Behavioral Research, 55*(3), 454-477. https://doi.org/10.1080/00273171.2019.1647822

Fandakova, Y., Selmeczy, D., Leckey, S., Grimm, K. J., Wendelken, C., Bunge, S. A., & Ghetti, S. (2017). Changes in ventromedial prefrontal and insular cortex support the development of metamemory from childhood into adolescence. *Proceedings of the National Academy of Sciences*, *114*(29), 7582-7587. https://www.pnas.org/doi/abs/10.1073/pnas.1703079114

Ferrer, E., & McArdle, J. J. (2004). An Experimental Analysis of Dynamic Hypotheses About Cognitive Abilities and Achievement From Childhood to Early Adulthood. *Developmental Psychology*, *40*, 935–952. https://doi.org/10.1037/0012-1649.40.6.935

Ferrer, E., & McArdle, J. J. (2010). Longitudinal Modeling of Developmental Changes in Psychological Research. *Current Directions in Psychological Science*, *19*(3), 149-154. https://doi.org/10.1177/0963721410370300

Ferrer, E., O'Hare, E., & Bunge, S. (2009). Fluid reasoning and the developing brain. *Frontiers in Neuroscience*, *3*, 3. https://doi.org/10. 3389/neuro.01.003.2009

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*, 466–491. https://doi.org/10. 1037/1082-989X.9.4.466

Ghisletta, P., & McArdle, J. J. (2012). Latent Curve Models and Latent Change Score Models Estimated in R. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(4), 651-682. https://doi.org/10.1080/10705511.2012.713275

Green, C. T., Bunge, S. A., Briones Chiongbian, V., Barrow, M., & Ferrer, E. (2017). Fluid reasoning predicts future mathematical performance among children and adolescents. *Journal of Experimental Child Psychology*, *157*, 125–143. https://doi.org/10.1016/j.jecp.2016.12.005

Hunter, M. D. (2018). State Space Modeling in an Open Source, Modular, Structural Equation Modeling Environment. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(2), 307–324. https://doi.org/10.1080/10705511.2017.1369354

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M. M., Moutoussis, M., Goodyer, I. M., Bullmore, E., Jones, P. B., Fonagy, P., Lindenberger, U., & Dolan, R. J. (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience, 33,* 99-117. https://doi.org/10.1016/j.dcn.2017.11.007

Lee, G., Arieli, R., Ryou, Y. J., & Martin, P. (2023). The bidirectional relationship between depressive symptoms and functional limitations among centenarian survivors in their 80s: Testing bivariate latent change score models. *Aging & Mental Health, 1-9. https://doi.org/10.1080/13607863.2023.2177830*

Liu, S., Wang, Y.-S., Zhang, Q., Zhou, Q., Cao, L.-Z., Jiang, C., Zhang, Z., Yang, N., Dong, Q., & Zuo, X.-N. (2021). Chinese Color Nest Project: An accelerated longitudinal brain-mind cohort. *Developmental Cognitive Neuroscience*, *52*, 101020. https://doi.org/10.1016/j.dcn.2021.101020

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analyses. *Structural equation modeling: Present and future: A Festschrift in honor of Karl Jöreskog* (pp. 7–46). Scientific Software International.

McArdle, J. J. (2009). Latent Variable Modeling of Differences and Changes with Longitudinal Data. *Annual Review of Psychology*, *60*(1), 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology*, *38*, 115–142. https://doi.org/10.1037/0012-1649.38.1.115

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In *New methods for the analysis of change* (pp. 139–175). American Psychological Association. https://doi.org/10.1037/10409-005

Power, M. J. E., Steptoe, A., Kee, F., & Lawlor, B. A. (2019). Loneliness and social engagement in older adults: A bivariate dual change score analysis. *Psychology and aging*, *34*(1), 152. https://doi.org/10.1037/pag0000287

Mistler, S. A., & Enders, C. K. (2012). Planned missing data designs for developmental research. In *Handbook of developmental research methods* (pp. 742–754). The Guilford Press.

Miyazaki, Y., & Raudenbush, S. W. (2000). Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychological Methods, 5*(1), 44-63. https://doi.org/10.1037/1082-989X.5.1.44

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended Structural Equation and Statistical Modeling. *Psychometrika*, *81*(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8

Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development.* Academic Press.

Oh, J., Chopik, W. J., Konrath, S., & Grimm, K. J. (2020). Longitudinal Changes in Empathy Across the Life Span in Six Samples of Human Development. *Social Psychological and Personality Science*, *11*(2), 244–253. https://doi.org/10.1177/1948550619849429

O'Rourke, H. P., Fine, K. L., Grimm, K. J., & MacKinnon, D. P. (2022). The importance of time metric precision when implementing bivariate latent change score models. *Multivariate Behavioral Research, 57(4), 561-580.* https://doi.org/10.1080/00273171.2021.1874261

Oud, J. H. L. (2001). Quasi-longitudinal Designs in SEM state Space Modeling. *Statistica Neerlandica*, *55*(2), 200-220. https://doi.org/10.1111/1467-9574.00165

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous Time Modeling of Panel Data by means of SEM. In K. van Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Longitudinal Research with Latent Variables* (pp. 201–244). Springer. https://doi.org/10.1007/978-3-642-11760-2_7

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of sem. *Psychometrika*, *65*(2), 199–215. https://doi.org/10.1007/BF02294374

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17,* 354–373. https://doi.org/10.1037/a0029315

Rhemtulla, M., & Hancock, G. R. (2016). Planned Missing Data Designs in Educational Psychology Research. *Educational Psychologist*, *51*(3–4), 305–316. https://doi.org/10.1080/00461520.2016.1208094

Rice, K. G., Montfort, A. K., Ray, M. E., Davis, D. E., & DeBlaere, C. (2018). A latent change score analysis of emotion regulation difficulties and evaluative threat in STEM. *Journal of Counseling Psychology*, *66*, 158–169. https://doi.org/10.1037/cou0000325

Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A Continuous-Time Approach to Intensive Longitudinal Data: What, Why, and How? In K. van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous Time Modeling in the Behavioral and Related Sciences* (pp. 27–54). Springer International Publishing. https://doi.org/10.1007/978-3-319-77219-6_2

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, *64*, 92–107. https://doi.org/10.1037/h0022371

Voelkle, M. C., & Oud, J. H. L. (2013). Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *British Journal of Mathematical and Statistical Psychology*, *66*(1), 103–126. https://doi.org/10.1111/j.2044- 8317.2012.02043.x

Voelkle, M. C., & Oud, J. H. L. (2015). Relating Latent Change Score and Continuous Time Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(3), 366–381. https://doi.org/10.1080/10705511.2014.935918

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*, 176–192. https://doi.org/10.1037/a0027543

Wendelken, C., Ferrer, E., Ghetti, S., Bailey, S. K., Cutting, L., & Bunge, S. A. (2017). Frontoparietal structural connectivity in childhood predicts development of functional connectivity and reasoning ability: A large-scale longitudinal investigation. *Journal of Neuroscience*, *37*(35), 8549-8558. https://www.jneurosci.org/content/37/35/8549.abstract

You, D., Hunter, M., Chen, M., & Chow, S.-M. (2020). A Diagnostic Procedure for Detecting Outliers in Linear State–Space Models. *Multivariate Behavioral Research*, *55*(2), 231–255. https://doi.org/10.1080/00273171.2019.1627659

Zainal, N. H., & Newman, M. G. (2021). Depression and executive functioning bidirectionally impair one another across 9 years: Evidence from within-person latent change and cross-lagged models. *European Psychiatry*, *64*(1), e43. https://doi.org/10.1192/j.eurpsy.2021.2217

Appendix

Figure A

Relative bias for correlation between latent variables

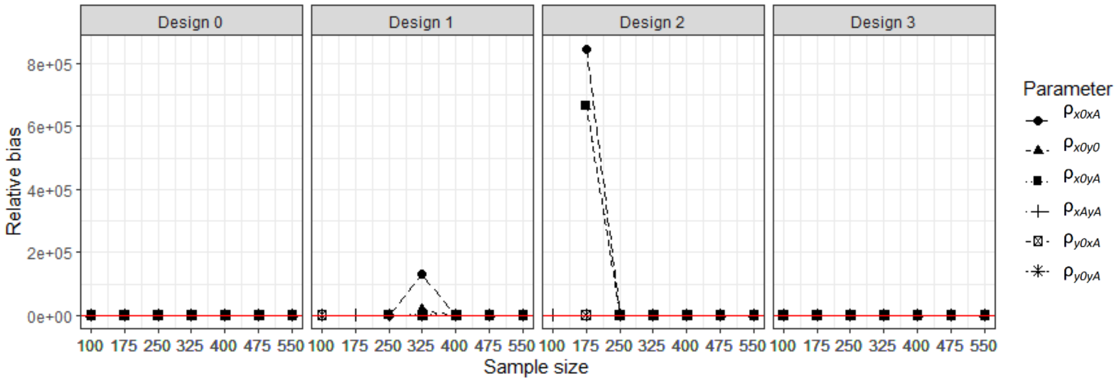
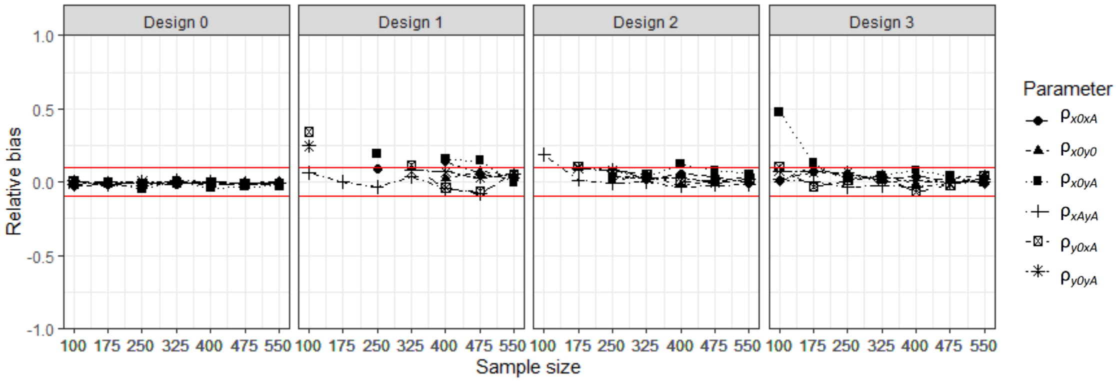


Figure B

Relative bias for correlation between latent variables without abnormal results



Note: An abnormal result was considered values >1 or < -1

As seen, abnormal values emerged across replicas that lead to improper relative bias values. Given this, we computed Median Relative Bias (MRB, Figure 8).